Volume 4 • Number 1 • 2001
VALUE IN HEALTH

US Regulation of Pharmaceutical Outcomes Research

Presented at the ISPOR 5th Annual International Meeting, Crystal City, VA, 24 May 2000

Laurie B. Burke, RPh, MPH

Chief, Evidence Review Branch, Division of Drug Marketing, Advertising and Communications (DDMAC), Food and Drug Administration (FDA)

Keywords: FDA, outcomes research, pharmaceuticals, regulation

Outcomes research is a subset of clinical evaluation that is increasingly being defined as a separate scientific discipline. It produces information that appears to be particularly attractive for use in pharmaceutical product promotion. As a result, FDA frequently must determine when this information is adequate to support labeling and advertising claims. Discussion of outcomes evidence is complicated by a lack of consistency in the terminology used, but perhaps more importantly, by a lack of understanding, and perhaps agreement on, the evidentiary standards that should apply to these data. FDA interacts with ISPOR and others to clarify issues and facilitate policy development in this area.

This lack of clarity is at least partly the result of the very different but overlapping categories of what appears to fall under the umbrella of outcomes research, according to the focus of ISPOR and others (Figure 1). Each category has its own special set of scientific and regulatory issues. For example, health-related quality of life (HRQL) is assessed in ways similar to other clinical endpoints but poses major problems because it usually assesses multiple endpoints so that attaining consistent results is difficult, and because the meaning of its results is often not straightforward. On the other hand, economic endpoints, although they can be measured as clinical endpoints in trials, are often assessed using observational methods and models, and raise special problems of generalizability.

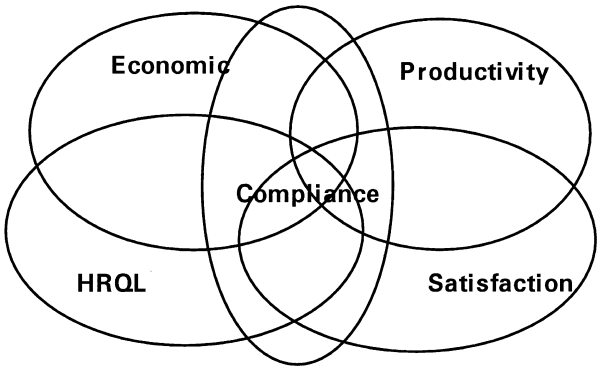

Figure 1 Categories of outcomes information.

Policy development for outcomes evidentiary requirements is complex and case-specific, dependent upon outcomes measurement, the disease, the intervention, and population studied. This is particularly evident in the area of HRQL outcomes research. Because the patient perspective is of great interest to audiences who make decisions about using drugs, pharmaceutical manufacturers commonly seek to assess that perspective in their studies and incorporate favorable findings in labeling and promotion in the form of such claims as "improves health-related quality of life," "maintains health-related quality of life," "contributes to patient well-being," "meaningful survival," and "expect a bright future." The first issue to be addressed when reviewing claims such as these is what meaning would be given to the claims by a recipient of the

Address correspondence to: Laurie B. Burke, Chief, Evidence Review Branch, DDMAC, U. S. Food and Drug Administration, 5600 Fishers Lane, Mail Stop HFD-42, Rockville, MD 20857. E-mail: burkel@cder.fda.gov

The views expressed in this article are those of the author and do not represent the official opinion or policy of the Food and Drug Administration.

 1098-3015/01/$15.00/5 5–7

information. Would the claim lead a patient or a physician to believe that the product will have the subjective effect described, including well-being, meaningful and bright? If so, the Federal Food Drug and Cosmetic Act and its regulations require that the claim, like any other, be backed by *substantial evidence* of that effect.

In some cases a claim is considered by FDA to be *puffery*, i.e., part of the way effects of a drug are described to make them attractive in promotion. Face validity can be adequate evidence for support for such statements. For example, "convenient" may be used to describe a product that is approved for twice daily dosing when compared to a product that requires four doses per day, as long as there is no implication of superior effectiveness. The line between puffery and claims that require substantiation can be difficult to draw, however, as can be seen in the above examples. In the case of "a bright future," if a product has been shown (with substantial evidence) to have a survival benefit, must the product also have been shown to increase the likelihood that the future will be bright? Or, just by reason of the fact that longer life can be expected can one assume the future to be bright? The answer to these questions depends on the disease or condition treated, the population indicated, and the drug.

If the statement represents a new claim, e.g., it actually uses the term HRQL, a more demanding review is in order. In addition to the usual concerns about design and analysis, there are special concerns with the assessment instrument. The instrument is scrutinized with all of its accompanying documentation regarding development and validation. We determine how instrument items, domains and scoring algorithms were generated, and whether validity, reliability, and internal consistency have been documented in a way that is appropriate for the study design and population. Documentation of the instrument's linearity of response, responsiveness to change and definitions of important change in score are then reviewed in light of the study protocol. An adequate study protocol will include details of the instrument administration, blinding, and plans for dealing with missing data. An adequate data analysis plan is a critical point of the review, particularly in light of the multiple comparisons almost invariably generated in HRQL studies. It is not unusual to see an analysis plan for the HRQL endpoints that is fully separate—perhaps even an appendix—from the primary efficacy endpoints. This makes the possibility of finding significant benefit remote, since only the most overwhelming results could be persuasive. FDA has heard that this situation results when the outcomes portion of the study design is developed in a different area of the company during product development. It would behoove the company to consolidate all study objectives and outcomes before completing the clinical studies. Moreover, developing an HRQL endpoint needs the same sort of attention as any other endpoint. Since it is not reasonable to imagine that an intervention would affect all HRQL measures similarly, it would be sensible to identify in advance the scales and subscales expected to respond.

Once the existence of an adequate protocol is determined, a thorough review of the study results and interpretation of those results is possible. The planned relationship between the HRQL endpoints and the primary efficacy outcome is critical since, without showing an effect on the primary efficacy outcome (i.e., without demonstrating that the drug works in the study population), it is doubtful that a HRQL claim can be supported.

Even in the case of adequate instrument development, validation, study protocol, data analysis and interpretation, FDA often questions whether the HRQL outcomes add to what we already know about a product's effects on symptoms and functional status. If the overall result is driven by the very effects already known for the drug, does the HRQL language add anything at all? For example, what is the added value of an outcome that combines symptoms and functional status into a single metric if we already know the drug's impact on the individual components? Isn't it in fact more informative to look at the symptoms? Further, apart from reference to the instrument, what words can be used to describe the results? Experts disagree about whether an effect on one domain of an instrument (e.g., physical symptoms) can support a claim if there is no impact on the overall HRQL index. To further complicate the issue, how can a claim such as "lead more active lives" be supported at all if no significant improvement is observed on a related objective measure (e.g., exercise tolerance). The HRQL scales that have been used to date in labeling, such as the Asthma Quality of Life Questionnaire and the Minnesota Living with Heart Failure scales, are primarily patient-derived physical disease severity scales, a kind of HRQL scale to be sure, but much closer to conventional symptom assessments.

Even though the multidimensional nature of HRQL is fairly well accepted by experts in this area of research, FDA has sometimes been presented with evidence from single-item questions to

support claims. In general, an effect on a single item from an instrument will not provide credible evidence of a benefit. A consistent effect on a defined set of items (i.e., a domain) over several trials, however, could be persuasive.

FDA has frequently objected to HRQL claims that are supported only by studies that have sought to demonstrate the impact or burden of disease on HRQL without any assessment of the benefit of the treatment. Companies have sought to imply that the treatment will reverse this negative disease impact. That sort of evidence cannot be expected to meet the statutory requirements for a claim.

Two adequate and well-controlled studies are generally required to support any clinical treatment outcome claim. The purpose of this requirement is to ensure that the evidence is strong enough to reach a conclusion and it not the result of chance or bias. It would be very difficult to make a persuasive case for a study that was not double blind. In addition, adequate validation of any measurement instrument (including HRQL) includes assessment of its measurement properties in a clinical trial setting before using the instrument in a trial to demonstrate a treatment effect. In many cases, the HRQL instrument is not adequately validated before a clinical study (e.g., a phase III study) is underway. In such cases, validation analyses can be designed and added to the data analysis plan before the study blind is broken. If the validation is adequate, a second clinical trial may serve as confirmatory evidence of a prospectively designated HRQL effect in the first trial if both studies are rigorously designed.

Because of the developing nature of this field, HRQL claims can be highly misleading unless they are accompanied by full disclosure of the meaning of the claim, the way the claim is supported, and the limitations of the claim. FDA is addressing HRQL issues on a case by case basis at the present time. As experience accumulates and the field advances, the Agency plans to issue guidance in this area. Until FDA guidance is available, requests for advice on development plans to support these outcomes are answered as time permits. If a company asks for our review of a study protocol, we can only review the plan in light of the company's intended labeling or advertising claims. Advice is then specific to the proposed claims and to the drug and condition reviewed.

Satisfaction and preference claims represent another group of outcomes claims for which policy is under development. FDA will object to a claim such as "satisfied patients" or "preferred treatment" unless it is supported by substantial evidence and its meaning is clear. Patient testimonials, single-arm studies, or surveys do not represent substantial evidence for these outcomes. Experience with these claims is growing, perhaps as a result of the need for managed care organizations to compete for patients. Study design principles that are specific to satisfaction are emerging. At present, FDA reviews these claims and their supporting evidence on a case by case basis.

Another burgeoning area of outcomes research and promotional claims review is productivity assessment. Developing interest in this type of claim in promotion seems to be related to the growing attention of employers to the cost of prescription drug benefits. Clinical trials sometimes incorporate productivity measures as secondary endpoints. When using a productivity measure as a treatment outcome, all the same concerns for instrument and study adequacy discussed above apply. Productivity can also be an economic term, and when used as an input to an economic model to produce cost estimates, standards that are appropriate for economic analyses would apply. The FDA Modernization Act of 1997 amended the Federal Food Drug and Cosmetic Act (FFDCA) by adding another evidence standard specific to economic information that is consistent with product approved indications and is disseminated exclusively to formulary committees or similar entities. As for other developing outcomes claims, FDA reviews economic claims and their supporting evidence on a case by case basis. The eventual development of guidance for such claims will depend on whether regulatory need and the FDA's experience justifies such guidance.

The future of outcomes research in drug development will be shaped by who the consumers of outcomes information are and the perceived added value of these outcome measures. It is likely that clinicians, patients and managed care will eventually understand, embrace and demand data on some of these outcome measures for use in their deliberative processes. In the meantime, development of this field of research can benefit by integration of outcomes researchers with the rest of the clinical development team to facilitate production of adequate evidence to meet developing regulatory hurdles.